

UNITED STATES PATENT APPLICATION

for

**JPEG 2000 FOR EFFICIENT IMAGING IN A CLIENT/SERVER  
ENVIRONMENT**

Inventors:

Martin Boliek

Gene K. Wu

Michael J. Gormish

prepared by:

BLAKELY, SOKOLOFF, TAYLOR & ZAFMAN LLP  
12400 Wilshire Boulevard  
Los Angeles, CA 90025-1026  
(408) 720-8598

File No.: 074451.P134

**EXPRESS MAIL CERTIFICATE OF MAILING**

"Express Mail" mailing label number: EL672750400US

Date of Deposit: 6/27/01

I hereby certify that I am causing this paper or fee to be deposited with the United States Postal Service "Express Mail Post Office to Addressee" service on the date indicated above and that this paper or fee has been addressed to the Assistant Commissioner for Patents, Washington, D. C. 20231

Mara E. Brown  
(Typed or printed name of person mailing paper or fee)

Mara E. Brown  
(Signature of person mailing paper or fee)

6/27/01  
(Date signed)

## JPEG 2000 FOR EFFICIENT IMAGING IN A CLIENT/SERVER ENVIRONMENT

### FIELD OF THE INVENTION

[0001] The present invention relates to the field of client/server systems ; more particularly, the present invention relates to client/server systems that provide for imaging a JPEG 2000 codestream.

[0002]

### **BACKGROUND OF THE INVENTION**

[0003] Digital image processing has undergone two significant evolutions in the 1990s. First, digital images are generally becoming higher resolution and higher quality. The consumer has access to near-photographic quality three megapixel digital cameras, for example. Scanners and printers of 600 dpi, or higher, are commonplace. Second, with the advent of the “populist” Internet and wireless technology, high speed networking now connects many types of heterogeneous display devices.

[0004] Image compression is vital as image size grows. However, because of the need to serve many different types of display devices with the same type of image, a new kind of image compression is required — image compression that is flexible at transmission and decode, not just encode. The JPEG 2000 image coding standard, ITU-T Rec.T.800/ISO/IEC 15441:2000

JPEG 2000 Image Coding System, allows one encoded image to be decoded at different resolutions, bit-rates, and with different regions without decoding any more than the minimum necessary amount of data.

[0005] The JPEG 2000 standard divides an image into tiles (rectangular regions), wavelet transform decompositions (different resolutions), codes coefficient bit-planes called layers (progressive bit-rate), components (e.g., R, G, B), and precincts (regions of wavelet coefficients). These tile-resolution-layers-component-precinct units are independently coded into JPEG 2000 packets. These packets can be identified and extracted from the codestream without decoding. This allows only the data required by a given display to be extracted and transmitted. For example, a monitor might require only 75 dpi data at a high quality while a printer might require 600 dpi data at a low quality. Both could be accommodated from one codestream.

[0006] The CREW image compression system served as a prototype for the features contained in the JPEG 2000 standard. For more information, see M. Boliek, M. J. Gormish, E. L. Schwartz, A. Keith, "Decoding compression with reversible embedded wavelets (CREW) codestreams," *Electronic Imaging*, Vol. 7, No. 3, July 1998. An earlier paper on JPEG 2000 codestream "parsing" shows how to access, extract, and rearrange the data in a JPEG

2000 codestream. For more information, see G.K. Wu, M. J. Gormish, M. Boliek, "New Compression Paradigms in JPEG2000," SPIE San Diego, July 2000. Another paper on JPEG 2000 codestream syntax shows how a codestream could be arranged for progressive transmission for a specific user interaction. For more information, see M. Boliek, J. S. Houchin, G. Wu, "JPEG 2000 Next Generation Image Compression System Features and Syntax," Int. Conf. On Image Processing 2000, Vancouver, Canada, 12 September 2000.

[0007] JPEG 2000 is a file format and compression algorithm. It does not specify the protocol or algorithms necessary to take advantage of the features in a client/server architecture, for example. The standard is similar to the Flashpix file format standardized by a consortium of companies now called the Digital Imaging Group. For more information, see Digital Imaging Group, "The Flashpix image format," with a world wide web site at [digitalimaging.org](http://digitalimaging.org). Using the original JPEG file format, Flashpix is not as efficient as JPEG 2000. However, unlike JPEG 2000, Flashpix is paired with a protocol for interacting with images over the Internet called the Internet Imaging Protocol (IIP). For more information, see Digital Imaging Group, "The Internet Imaging Protocol."

## SUMMARY OF THE INVENTION

[0008] In one embodiment, the system comprises a server and a client. The server stores a compressed codestream corresponding to image data. The client is coupled to the server via a network environment. The client includes a memory having an application and a data structure stored therein. The data structure identifies positions of packets of the compressed codestream on the server and identifies data of the compressed codestream already buffered at the client.

[0009] In one embodiment, the client requests bytes of the compressed codestream from the server that are not already stored in the memory and generates decoded image data requested by a user from the bytes of the compressed codestream requested from the server and any portion of the compressed codestream previously stored in the memory necessary to create the image data.

## BRIEF DESCRIPTION OF THE DRAWINGS

[0010] The present invention will be understood more fully from the detailed description given below and from the accompanying drawings of various embodiments of the invention, which, however, should not be taken to limit the invention to the specific embodiments, but are for explanation and understanding only.

[0011] **Figure 1** illustrates packets arranged in a JPEG 2000 codestream (pixel fidelity order on left, resolution order on right).

[0012] **Figure 2** illustrates an exemplary network environment.

[0013] **Figure 3** illustrates a flow diagram of a process for processing a JPEG 2000 codestream.

[0014] **Figure 4** illustrates a JP2 file with the length of the boxes preceding the codestream and the length of the main header denoted.

[0015] **Figure 5** illustrates the location of the packets in a JPEG 2000 codestream.

[0016] **Figure 6** is a block diagram of one embodiment of a computer system.

## DETAILED DESCRIPTION OF THE PRESENT INVENTION

[0017] A novel and useful system is described herein that includes a number of techniques for processing a JPEG 2000 or similar codestream in a client/server system. These include a data structure that tracks the length and location of every packet in the codestream both on the server and received by the client. In one embodiment, the JPEG 2000 TLM and PLM marker segments are used. Also included are interactive client/server protocol with clients making byte requests, and modifications of the Internet Image Protocol (IIP, owned by the Digital Imaging Group, a commercial consortium) for JPEG 2000.

[0018] The JPEG 2000 image compression system offers significant opportunity to improve imaging over the Internet. The JPEG 2000 standard is ideally suited to the client/server architecture of the web. With only one compressed version stored, a server can transmit an image with the resolution, quality, size, and region custom specified by an individual client. It can also serve an interactive zoom and pan client application. All of these can be achieved without decoding at the server while using only reduced, and potentially minimal, server computation, storage, and bandwidth.

[0019] The following description discusses some of the system issues involved in Internet imaging with JPEG 2000. The choices of the client, passing of control information, and the methods a server could use to serve the client requests are described herein. These issues include standard use of JPEG 2000 encoding and the decoding options.

[0020] In the following description, numerous details are set forth to provide a thorough understanding of the present invention. It will be apparent, however, to one skilled in the art, that the present invention may be practiced without these specific details. In other instances, well-known structures and devices are shown in block diagram form, rather than in detail, in order to avoid obscuring the present invention.

[0021] Some portions of the detailed descriptions which follow are presented in terms of algorithms and symbolic representations of operations on data bits within a computer memory. These algorithmic descriptions and representations are the means used by those skilled in the data processing arts to most effectively convey the substance of their work to others skilled in the art. An algorithm is here, and generally, conceived to be a self-consistent sequence of steps leading to a desired result. The steps are those requiring physical manipulations of physical quantities. Usually, though

not necessarily, these quantities take the form of electrical or magnetic signals capable of being stored, transferred, combined, compared, and otherwise manipulated. It has proven convenient at times, principally for reasons of common usage, to refer to these signals as bits, values, elements, symbols, characters, terms, numbers, or the like.

[0022] It should be borne in mind, however, that all of these and similar terms are to be associated with the appropriate physical quantities and are merely convenient labels applied to these quantities. Unless specifically stated otherwise as apparent from the following discussion, it is appreciated that throughout the description, discussions utilizing terms such as "processing" or "computing" or "calculating" or "determining" or "displaying" or the like, refer to the action and processes of a computer system, or similar electronic computing device, that manipulates and transforms data represented as physical (electronic) quantities within the computer system's registers and memories into other data similarly represented as physical quantities within the computer system memories or registers or other such information storage, transmission or display devices.

[0023] The present invention also relates to apparatus for performing the operations herein. This apparatus may be specially constructed for the

required purposes, or it may comprise a general purpose computer selectively activated or reconfigured by a computer program stored in the computer. Such a computer program may be stored in a computer readable storage medium, such as, but not limited to, any type of disk including floppy disks, optical disks, CD-ROMs, and magnetic-optical disks, read-only memories (ROMs), random access memories (RAMs), EPROMs, EEPROMs, magnetic or optical cards, or any type of media suitable for storing electronic instructions, and each coupled to a computer system bus.

[0024] The algorithms and displays presented herein are not inherently related to any particular computer or other apparatus. Various general purpose systems may be used with programs in accordance with the teachings herein, or it may prove convenient to construct more specialized apparatus to perform the required method steps. The required structure for a variety of these systems will appear from the description below. In addition, the present invention is not described with reference to any particular programming language. It will be appreciated that a variety of programming languages may be used to implement the teachings of the invention as described herein.

[0025] A machine-readable medium includes any mechanism for storing or transmitting information in a form readable by a machine (e.g., a computer). For example, a machine-readable medium includes read only memory ("ROM"); random access memory ("RAM"); magnetic disk storage media; optical storage media; flash memory devices; electrical, optical, acoustical or other form of propagated signals (e.g., carrier waves, infrared signals, digital signals, etc.); etc.

#### **[0026] Overview**

[0027] The following description sets forth protocols and algorithms for interacting over a client/server network in processing a JPEG 2000 or similar codestream. A description of a networked environment is set forth, followed by descriptions of different environments where the computational burden is distributed differently between the client and server machines. In the first environment (referred to herein Smart Client, Challenged Server), the burden of computation is weighted toward the client. This is good for networks where the client is a powerful capable machine like a personal computer (PC) or a laptop. The second environment (referred to herein Smart Server, Challenged Client) offloads more of the computational burden

from the client to the server. This is useful for less powerful machines such as cell phones, PDAs, etc. Note that a server could be designed to be capable of either model depending on the client's capability. Thus, heterogenous client devices can share a network and achieve just the correct level of server support.

#### **[0028] Division of a JPEG 2000 Image**

[0029] In JPEG 2000, a typical image consists one or more components (e.g. red, green, blue). Components are rectangular arrays of samples. These arrays are further divided into regular rectangular tiles. On a tile by tile basis the components can be decorrelated with a color space transformation. After color space transformation, every tile-component is compressed independently.

[0030] After each tile-component is transformed with a wavelet transformation. The multiple scale characteristic of the wavelet transformation provides groupings of coefficient data (LL subband, and HL, LH, HH subbands) capable of reconstructing the tile-component at different resolutions.

[0031] The coefficient subbands are further grouped into code-blocks (regular rectangular regions covering each subband). The bit planes, starting with by the most significant bit of each coefficient in the code-block, are coded using a context model and arithmetic coder. The coding produces several coding passes (up to three per bit plane) in order from the most significant to the least significant.

[0032] After all the coefficient data is coded the coding passes are arranged in packets. Each packet represents a collection of coding passes from some, or all of the code-blocks, at a given resolution and precinct, in a tile-component. In other words, a packet provides one unit of refinement for a given resolution within a tile-component. For each resolution in each tile-component, there is a strict order for the packets. However, the packets from different resolutions and tile-components can be interleaved in a variety of ways. The packet is the minimum unit of coded data that is easily accessible in the codestream.

[0033] Figure 1 shows two examples of packets orders. On the left, the packet order is in pixel fidelity order (3 resolutions, 3 layers, one precinct, one tile-component). On the right, the packet order is resolution order (A, D, G, B, E, H, C, F, I).

[0034] The syntax of the codestream follows the division of the image. There is a main header at the beginning of the codestream. This header contains markers that describe the image characteristics and the coding style and other parameters that apply to the whole image or individual components. Each tile-part has a header. These tile-parts are indexed to indicate order. The first tile-part header of a tile contains information that applies to whole tile or individual tile-components. The remaining tile-part headers include only order and length information for that tile-part and/or succeeding tile-parts.

#### [0035] An Exemplary Network

[0036] Figure 2 is a block diagram of one embodiment of a network environment 201 that may be used with the techniques described herein. In one embodiment, a server computer system 200 is coupled to a wide-area network 210. Wide-area network 210 may include the Internet or other proprietary networks including, but not limited to, America On-Line<sup>TM</sup>, CompuServe<sup>TM</sup>, Microsoft Network<sup>TM</sup>, and Prodigy<sup>TM</sup>. Wide-area network 210 may include conventional network backbones, long-haul telephone lines, Internet and/or Intranet service providers, various levels of network

routers, and other conventional mechanisms for routing data between computers. Using network protocols, server 200 may communicate through wide-area network 210 to client computer systems 220, 230, 240, which are possibly connected through wide-area network 210 in various ways or directly connected to server 200. For example, client 240 is connected directly to wide-area network 210 through direct or dial-up telephone or other network transmission line. Client 240 may be connected to wide-area network 210 via a wireless connection.

[0037] Alternatively, clients 230 may be connected through wide-area network 210 using a modem pool 214. Modem pool 214 allows multiple client systems to connect with a smaller set of modems in modem pool 214 for connection through wide-area network 210. Clients 231 may also be connected directly to server 200 or be coupled to server through modem 215. In another alternative network typology, wide-area network 210 is connected to a gateway computer 212. Gateway computer 212 is used to route data to clients 220 through a local area network 216. In this manner, clients 220 can communicate with each other through local area network (LAN) 216 or with server 200 through gateway 212 and wide-area network

210. Alternatively, LAN 217 may be directly connected to server 200 and clients 221 may be connected through LAN 217.

[0038] Using one of a variety of network connection mechanisms, server computer 200 can communicate with client computers 250. In one embodiment, a server computer 200 may operate as a web server if the World-Wide Web ("WWW") portion of the Internet is used for wide area network 210. Using the HTTP protocol and the HTML coding language, or XML, such a web server may communicate across the World-Wide Web with a client. In this configuration, the client uses a client application program known as a web browser such as the Netscape™ Navigator™, the Internet Explorer™, the user interface of America On-Line™, or the web browser or HTML translator of any other conventional supplier. Using such browsers and the World Wide Web, clients 250 may access graphical and textual data or video, audio, or tactile data provided by the web server 200.

**[0039] Smart Client, "Challenged" Server**

[0040] In one embodiment, a client represents a smart terminal, such as a personal computer, that provides requests to a server to obtain some amount of data corresponding to an image. The data being requested is part

of a codestream, such as a JPEG 2000 codestream, stored as a file at the server. The server receives the request for bytes for a particular file and transmits them to the client.

[0041] Figure 3 is one embodiment of a process performed by processing logic of the client (e.g., the client side application) to display an image requested by the user. With this process, a client side application is responsible for determining what data is needed and asking for the data from the server. The processing logic may comprise hardware (e.g., circuitry, dedicated logic, etc.), software (such as is run on a general purpose computer system or a dedicated machine), or a combination of both.

[0042] Referring to Figure 3, processing logic determines the image characteristics that the user requests (processing block 301). These may include region, resolution, precinct and/or quality.

[0043] Next, processing logic selects the data of the JPEG 2000 codestream that corresponds to these image characteristics (processing block 302), and determines what byte requests are necessary to receive this data based on what is already buffered at the client (processing block 303). The client determines which packets it needs and which packets it already has in order to generate a request for the packets it still needs. The client initially begins

with no packets and then as more requests are made, and as described in further detail below, the client retains packets of the codestream and stores them in a manner that provides the client easy access to the previously requested and obtained packets. In one embodiment, the client only retains packets for one type of interactive session. In one embodiment, this is performed using a data structure described below. However, this could be performed in a number of ways.

[0044] Using this information, processing logic issues byte range requests to the server (processing block 304). In one embodiment, the client specifies the data of the JPEG 2000 codestream that is needed by sending the starting point of the memory location at which the data is stored and the range of the amount of data that is requested. In an alternative embodiment, the starting and ending points of the memory locations storing the desired data are sent in the request.

[0045] In one embodiment, the user may use HTTP or a similar mechanism to obtain portions of the codestream which it desires. For example, the client may use POST or GET operations to request the information from the server. In an alternative embodiment, the client may specify the starting location (e.g., byte) and length or starting location (e.g., byte) and ending location

(e.g., byte) (or multiple sequences of these) in a universal resource locator (URL).

[0046] Processing logic integrates the received data, which is sent in encoded format, with the previously buffered data to create a correct JPEG 2000 codestream (processing block 305). In one embodiment, the packets are put in the order they appear in the original codestream.

[0047] The marker segments in the JPEG 2000 codestream may be changed to create a legal codestream that a JPEG 2000 compliant decoder may be able to handle. More specifically, the markers associated with the original JPEG codestream stored on the server indicate a certain number of tiles and tile parts, etc. These markers are forwarded with the requested bytes of the codestream. (The server does not necessarily know that it is providing a JPEG 2000 file; it simply receives requests and sends bytes.) However, since only a portion of the codestream may have been requested, the client modifies the markers so that the markers are correct for the codestream that is generated as a result of the integration process. Thus, the client creates a correct, or legal, codestream. For example, when a thumbnail (lower resolution) version of a codestream is requested, the server provides the requested data. However, the PLM values provided in the main header are

no longer correct. Since some packets which belong to the higher resolution are not included in the new codestream, the PLM values must be updated. Similarly, the Psot value (length from beginning of the first byte of the SOT (Start of tile-part) marker segment of the tile-part to the end of the data of that tile-part) of the SOT marker as well as the Ttlm and Ptllm values of the TLM marker must be updated to reflect the change.

[0048] Then processing logic decodes the newly generated JPEG 2000 codestream using a compliant JPEG 2000 decoder (processing block 306) and draws the image to the monitor (processing block 307).

[0049] The server services the request; however, the server need not have any special software for handling JPEG 2000 files. However, the control on the server (e.g., HTML or other control language) includes not only the file handle but a length of the main header as well. In response to the request, the server provides information to the client, including the requested bytes.

In one embodiment, the server is able to serve byte length requests. This could be accomplished with a number of methods. In one embodiment, a Common Gateway Interface (CGI) script is used to extract a consecutive portion of a file and transmit it to the client. The CGI script determines the bytes to send, creates TCI/IP packets, and sends the created packets.

[0050] The client creates a legal JPEG 2000 codestream from the transmitted data packets and decodes them using a generic decoder. To do so, the client assembles the portion of the codestream it receives into a complete legal JPEG 2000 codestreams for the decoder. The decoder may comprise a generic software decoder. The generic decoder expects to have access to the entire codestream (i.e., a legal codestream). In an alternative embodiment, the decoder is able to process different portions of a codestream by skipping around memory. A stream filter may be included to filter the stream prior to it being decoded by the decoder.

[0051] In an alternative embodiment, the decoder is an enhanced decoder that is able to handle codestreams that are not in order and that may be spread out in storage. In such a case, the headers may indicate the location of each of the necessary parts spread out through the memory, preferably through the use of pointers.

[0052] In still another embodiment, a restricted decoder is used, such as a decoder that can understand only one of the JPEG 2000 progression orders. In such a case, the data that is requested and received from the server may have to be reordered before being sent to the decoder.

[0053] With respect to JPEG 2000, the JPEG 2000 file may have fully specified TLM and PLM marker segments. With these marker segments, the length and starting point of every packet in the codestream is known after the main header is received by the client. In an alternative embodiment, this information is discerned without these marker segments by reading every tile-part header and possibly every packet header — a procedure that would be less efficient.

[0054] In one embodiment, the image and the JPEG 2000 codestream include the following characteristics. The image size, tile size, tile-parts, number of resolutions, and the layering pattern are reasonable. That is, the image is large enough that interaction makes sense versus transmission of the entire image. For example, in one embodiment, a 1024 x 1024 pixel color image or larger is used. The tile size is large enough to preserve image quality but small enough to allow selection of small portions of the file, e.g., 256 x 256 is reasonable. There are not too many tile-parts, because too many tile-parts increases the size of the TLM marker segment. If the TLM is adding more than 5% to the image size, it is probably inefficient. (When too many tile-parts occurs, it is easier to download the entire image rather than pass the signaling data.) The number of resolutions allows small enough images to

be extracted. For example, 4 to 6 decomposition levels might be reasonable to allow small enough images to be extracted. Also, the layering pattern includes a predictable, relatively equal distortion across the image. Preferably, it should also be fine enough to allow a good level of control over the rate yet not so fine as to create too many small packets. Small packets may not be the most efficient for access to the server and the number of packets increases the size of the PLM marker segment.

[0055] The client/server interaction may be more efficient if the JPEG 2000 files are originally encoded according to one or more of the suggestions below. In one embodiment, the file only contains one progression ordering (e.g., resolution level-layer-component-position), one partition, and COD, COC, QCD, and QCC marker segments only in the main header. The tile-part headers only contain SOT and SOS marker segments. The components of the image are all treated the same for the examples described herein. In theory, none of these assumptions are required for extracting the correct parts of the codestream. However, in practice these are useful for efficiency.

Please note that that this organization of a JPEG 2000 codestream is quite common. Also note that no special use of the file format features of JPEG

2000 codestream is discussed herein as such would be well-known to one skilled in the art.

**[0056] Retrieving the Information in the Main Header of the Codestream**

**[0057]** For a numerical example, the well-known Lena image is used with three components, four tiles (one tile-part per tile), three resolutions levels, and two layers. There are 72 packets ( $3 \times 4 \times 3 \times 2 = 72$ ). The browser application receives the filename and a length from the HTML source (e.g., a web page with the embedded image or a page with a link to the image). An XML source could be used as well. For example, in one embodiment, the format is the following:

```

```

where the headerlength (309 in this example) corresponds to the data in the file format boxes in front of the codestream and the syntax main header of the codestream. Figure 4 shows this file and the headerlength. Referring to Figure 4, the JP2 file 400 includes file format boxes 401 and a codestream box 402. Codestream box 402 includes a main header 403 and tile-part headers and data 404. The headerlength 405 is equal to file format boxes 401 and main header 403. Note that Figure 4 is not to scale.

[0058] The client makes a request to the server for the data corresponding the header length. In one embodiment, the request causes the server to call a CGI script on the server that streams only the bytes between the range requested. For example, syntax of this request could be the following:

```
cgiByteRequest:lena.jp2:0:309
```

[0059] The client reads the information in the file format boxes and uses it as necessary. In the case of a JPEG 2000 codestream, the main header of the codestream box includes two key marker segments useful for further interaction, the TLM and the PLM marker segments, in addition to the coding and quantization marker segments. Together these two marker segments provide a byte map to every packet. The packets are distinguished by tile, component, resolution, and layer.

[0060] For example, if an image has three components, is divided into four tiles (with one tile-part per tile), has three resolution levels (two wavelet decompositions), two layers, and only one partition per tile, then this image would have 72 packets and a packet represents one partition of one layer of one resolution of one tile-component. In this case, the TLM has the following values:

```
TLM->marker =      0xFF55      // marker number  
TLM->Ltlm =        24         // length of the marker segment in bytes
```

```

TLM->Ztlm = 0           // index of this TLM marker segment
TLM->Stlm = 0x50        // setting for Ttlm as 8 bits and Ptlm as 32
                         //bits
TLM->Ttlm0 = 0          // tile 0
TLM->Ptlm0 = 80,996     // tile 0 length
TLM->Ttlm1 = 1
TLM->Ptlm1 = 74,474
TLM->Ttlm2 = 2
TLM->Ptlm2 = 90,320
TLM->Ttlm3 = 3
TLM->Ptlm3 = 70,296

```

[0061] The PLM marker segment describes the length of the 72 packets as follows

```

PLM->marker = 0xFF57    // marker number
PLM->Lplm = 157         // length of the marker segment in bytes
PLM->Zplm = 0           // index of this PLM marker segment
PLM->Nplm0 = 38        // number of bytes for the first tile-part in the
                         //codestream, in this case tile 0
PLM->Iplm0,0 = 1895      // packet length
PLM->Iplm0,1 = 1802
...
PLM->Iplm0,17 = 16438
...
PLM->Nplm3 = 37        // number of bytes for the fourth tile-part in the
                         //codestream, in this case tile 3
PLM->Iplm3,0 = 1994      // packet length
PLM->Iplm3,1 = 1853
...
PLM->Iplm3,17 = 18,031

```

[0062] The starting points of each packet can be discerned with the information now available at the client. The starting point of the first tile is

known from the headerlength value in the script that calls out the image.

From that, the location of all the tile-parts are known. Figure 5 shows an example with two tile-parts and two packets per tile-part. Referring to Figure 5, packets 501 are packets associated with tile-part headers 502 and packets 503 are packets associated with tile-part header 504.

[0063] In one embodiment, the tile-part headers only have SOT and SOS marker segments. Thus, the start of the first packet is 12 bytes after the start of the tile-part. (If the length of the tile-part header was not known it could be deduced by subtracting the sum of the packet lengths in the tile-part from the length of the tile-part.)

[0064] From this information, and the known order of the packets, the exact location and length of each packet is known. The client can create a data structure that lists the locations of all the packets on the server side and relates that to data that has been received on the client side. For the above example, the data structure might contain the data in Table 1. Note that this information can be generated regardless of the number of tile-parts for a given tile or the order in the codestream.

[0065] Table 1 — Data Structure to Determine the Position of Packets on Both the Server and the Client

Tile	Resolution	Layer	Component	Precinct	Server start offset	Packet length	Client start offset
0	0	0	0	0	323	1895	null
0	0	0	1	0	2218	1802	null
0	0	0	2	0	4020	1658	null
0	0	1	0	0	5678	1608	null
...							
3	2	1	0	0	274,063	9811	null
3	2	1	1	0	283,874	14,490	null
3	2	1	2	0	298,364	18,031	null

[0066] In one embodiment, multiple client applications requesting data from the server utilize the table and the image data is stored as indicated by the table. In such a case, the client may include the cache manager that manages the JPEG 2000 images and any requests by a web browser for such an image is first checked in the table to see if it is present. Thus, the table may be a shared resource shared between a number of applications.

[0067] *Determining the Data to Request*

[0068] In one embodiment, the three main image requests that the user can control are region (which tiles), size (resolution), and quality or speed (layers). The user may indicate to the client the portion of a particular image

they wish to view through a user interface. In response to the indication, the client generates a request to the server to obtain the portion of the codestream that it does not have that corresponds to the requested image data. The user interface allows an image to be selected in a particular size, bit rate, or region to be chosen.

[0069] In one embodiment, the user interface may comprise a web site with a series of thumbnails from which the user may select with a cursor control device. In such a case, the user causes the request to a server by selecting one or more of the images (e.g., clicking on one or more of the images) to obtain an image of greater size. The selection is received by a browser application on the client, which generates the request.

[0070] In an alternative embodiment, the client may utilize a double click to specify their request. In still another embodiment, a box may appear on the screen that requests vertical and horizontal dimensions or DPI that a user may indicate a portion of an image that they wish to view. In yet another embodiment, a slide bar may appear on the screen. The slide bar may be a quality slide bar that allows the user to specify the quality of the image that is to be displayed. In another embodiment, a set of buttons may be used to make the request. In another embodiment, the user draws a rectangle on the

screen to specify the region of an image they wish to view at a different level of resolution. In still another embodiment, an algorithm may be used to find a particular location on a page, such as, by searching for particular text or a graphic that appears in the image. In still another embodiment, simply rolling a mouse over and highlighting a portion of the image causes the client to generate a request associated with that highlighted region.

[0071] In one embodiment, no user interface is utilized at all. For example, inside an HTML page, a designer may have predetermined the size of image that is going to be supplied to the user. Therefore, when the user loads a page containing an image the size of the image is at a specified or preset rate.

**[0072] *Data Request, Receiving, and Merging***

[0073] In one embodiment, a CGI script is used to process the request. First, the correct packets are requested. Next, the client receives packets and merges streams. That is, the client combines the packets received from the server in response to the request with those previously stored on the client. This is performed using the Client start offset column of Table 1.

[0074] The processing continuously changes the data structure as the new data is received or near in time thereafter. In this way, the data structure provides an accurate indication of which portions of a JPEG 2000 codestream are buffered by the client. Thereafter, when a second request occurs, the new data is merged.

**[0075] Smart Server, "Challenged" Client**

[0076] The Internet Imaging Protocol (IIP) provides a definition of client server communication for the exchange of image data. It was primarily designed to take advantage of the Flashpix file format in a networked environment. Since Flashpix was designed to store multiple resolutions of an image using compressed 64x64 tiles, it may be adapted to use with a JPEG 2000 codestream.

[0077] In fact, the IIP could be used to serve JPEG 2000 images without any changes whatsoever to the protocol or the client, if the server converts all requested tiles to the JPEG DCT compression method as they were requested. This allows the server to realize the compression savings provided by JPEG 2000's ability to store multiple resolutions in one file, but it would not provide any transmission time savings and suffers from

processing time to recompress portions of the image and suffers from reduced quality due to JPEG compression. In such a case, the client has the capability to decode JPEG DCT compression.

[0078] A more useful approach is to amend the IIP specification to allow JPEG 2000 as a native compression type. This is a trivial modification as far as the protocol is concerned. Unfortunately, because Flashpix uses fixed size tiles (64 by 64) which are the same size on every resolution level, and JPEG 2000 uses arbitrary sized tiles which change with each resolution level (and thus maintain the same number of tiles on each resolution level), substantial changes to both the client and the server would be necessary to implement the new compression type.

[0079] An example of an IIP exchange is as follows from Annex 1 of Version 1.0.5 of the specification:

The client provides the following:

```
FIF=Moon.fpx&OBJ=IIP,1.0&OBJ=IIP,1.0& //client specifies user of IIP  
//protocol version 1.  
OBJ=Basic-info&OBJ=Comp-group,2,*&OBJ=Title //client requests  
//basic info about  
//compression Group  
//2 in the codestream
```

In response, the server provides:

```
IIP:1.0CRLF //indicates server is using IIP 1.0  
IIP-server:0.0CRLF //server indicated its server 0.0  
Max-size:1000 1000CRLF //specifies maximum resolution
```

Resolution-number:5CRLF	//specifies resolution number of //image as part of basic info
Colorspace,0-4,0:0 0 3 3 0 1 2CRLF	//specifies color space of image as //part of basic info
ROI:0 0 1.5 1.CRLF	//specifies region of intent of //interest
Affine-transform:0.86 -0.49 0 0.35 0.49 0.86 0 -0.3 0 0 1 0 0 0 0 1 CRLF	//specifies transform
Aspect-ratio:1.5CRLF	//specifies aspect ratio
Error/19:3 3 Filtering-valueCRLF	//specifies no filtering
Error/15:3 3 Color-twistCRLF	//specifies no color twist
Error/19:3 3 Contrast-adjustCRLF	// specifies no color adjust
Comp-group,2,0/785:dataCRFL	//server sending back //compression group 2 + 785 bytes //of data
Title/38:the moon in the skyCRLF	//server sends back title

Then the client provides:

FIF=Moon.fpx&OBJ=IIP,1.0&TIL=2,44&TIL=3,0-1	//client specifies that //it wants tile at //resolution 2, tile //44; and at //resolution 3, tiles 0- //1.
---	---

and the server responds by sending the requested data:

IIP:1.0CRLF	
Tile,2,44,0/12296:dataCRLF	//server sends tile at resolution 2, //its tile number is 44 and there are //12296 bytes of data.

Tile,3,0,0/980:dataCRLF

Tile,3,1,0/1011:dataCRLF

[0080] As described above, the first request from the client asks for some fundamental information about the flashpix file. The server responds with several items including an indication that the "Filtering-value," "Color-

twist," and "Contrast-adjust" are unavailable. Importantly, the client learns the image size and the maximum number of resolutions. The client then requests one tile at resolution 2 and two tiles (0 and 1) at resolution 3. In one embodiment, the server provides an initial HTML, XML or similar file, and then starts the IIP process. From this file, the client could ascertain the image size and maximum number or resolutions. In this manner, one round trip communication between the server and the client could be avoided.

[0081] Note that because Flashpix stores resolutions independently, there is no need to track transmission of lower resolution portions of an image. In one embodiment, a data structure at the server similar to Table 1 is used in which the server stores an indication of what portions of a JPEG 2000 codestream are stored at the client.

[0082] In one embodiment, with JPEG 2000, the interaction might be very similar. First, a client asks for fundamental information. If that information is included in the JPEG 2000 main header (or even just the SIZ tag), the client then determines which tiles to request. Unlike Flashpix, the same tiles may be requested for the same portion of the image regardless of the resolution. Note that for some JPEG 2000 files, the maximum resolution may not be defined because the number of resolutions can vary on a tile-by-tile basis;

however, in one embodiment, for purposes herein the system is limited to codestreams which do not contain COD or COC makers in the tile part headers as may be defined as a “profile” for some applications.

[0083] Given a request for Tile 0, at resolution 2, the server parses the JPEG 2000 codestream and locates all packets relevant to the request. These are all packets for the requested tile at a resolution less than or equal to the resolution requested and all layers. The definition of the tile object returned by a IIP server when using JPEG 2000 could be either all resolutions up to and including the named resolution or only the named resolution. In the second case, the following exchange might occur:

From the client:

FIF=Moon.fpx&OBJ=IIP,1.0&TIL=2,4

From the server:

IIP:1.0CRLF

Tile,0,4,0/597:dataCRLF

Tile,1,4,0/2296:dataCRLF

Tile,2,4,0/6296:dataCRLF

[0084] Of course, using IIP in this way limits the selection of portions of the bitstream to only resolution and spatial region. In one embodiment, the tile request syntax is:

TIL=res,tile[,sub]

[0085] For full use of JPEG 2000, this syntax could be modified to

TIL=res,comp,lay,prec,tile

where res is the resolution as currently defined, comp is the component, lay is the JPEG 2000 layer, and prec is the JPEG 2000 precinct. Example ranges include (0-2), or wildcards “\*,” could be used for any of the parameters.

[0086] Thus a client request might be the following:

=>  
TIL=0-2,\*,0-1,\*,5

to obtain all components and all precincts for the first 3 resolutions of the 5th tile, but only get the first 2 layers which could reduce the required bitrate substantially.

#### [0087] An Exemplary Computer System

[0088] Figure 6 is a block diagram of an exemplary computer system that may perform one or more of the operations described herein. Referring to Figure 6, computer system 600 may comprise an exemplary client 650 or server 600 computer system. Computer system 600 comprises a communication mechanism or bus 611 for communicating information, and a processor 612 coupled with bus 611 for processing information. Processor

612 includes a microprocessor, but is not limited to a microprocessor, such as, for example, Pentium™, PowerPC™, Alpha™, etc.

[0089] System 600 further comprises a random access memory (RAM), or other dynamic storage device 604 (referred to as main memory) coupled to bus 611 for storing information and instructions to be executed by processor 612. Such instructions may be those which when executed by processor 612 cause the operation of the client and/or server to be performed. Main memory 604 also may be used for storing temporary variables or other intermediate information during execution of instructions by processor 612.

[0090] Computer system 600 also comprises a read only memory (ROM) and/or other static storage device 606 coupled to bus 611 for storing static information and instructions for processor 612, and a data storage device 607, such as a magnetic disk or optical disk and its corresponding disk drive. Data storage device 607 is coupled to bus 611 for storing information and instructions.

[0091] Computer system 600 may further be coupled to a display device 621, such as a cathode ray tube (CRT) or liquid crystal display (LCD), coupled to bus 611 for displaying information to a computer user. An alphanumeric input device 622, including alphanumeric and other keys, may also be

coupled to bus 611 for communicating information and command selections to processor 612. An additional user input device is cursor control 623, such as a mouse, trackball, trackpad, stylus, or cursor direction keys, coupled to bus 611 for communicating direction information and command selections to processor 612, and for controlling cursor movement on display 621.

[0092] Another device that may be coupled to bus 611 is hard copy device 624, which may be used for printing instructions, data, or other information on a medium such as paper, film, or similar types of media. Furthermore, a sound recording and playback device, such as a speaker and/or microphone may optionally be coupled to bus 611 for audio interfacing with computer system 600. Another device that may be coupled to bus 611 is a wired/wireless communication capability 625 to communication to a phone or handheld palm device.

[0093] Note that any or all of the components of system 600 and associated hardware may be used in the present invention. However, it can be appreciated that other configurations of the computer system may include some or all of the devices.

[0094] Whereas many alterations and modifications of the present invention will no doubt become apparent to a person of ordinary skill in the art after

having read the foregoing description, it is to be understood that any particular embodiment shown and described by way of illustration is in no way intended to be considered limiting. Therefore, references to details of various embodiments are not intended to limit the scope of the claims which in themselves recite only those features regarded as essential to the invention.